Voronoi Cell: New Method for Allocation of Space among Atoms: Elimination of Avoidable Errors in Calculation of Atomic Volume and Density

A. GOEDE, R. PREISSNER, C. FRÖMMEL

Humboldt University, Institute of Biochemistry, Monbijoustrasse 2A, Berlin D-10117, Germany

Received 17 June 1996; accepted 19 November 1996

ABSTRACT: In computing the volume occupied by atoms and the density in proteins, one is faced with the problem of intersecting spheres. To estimate either, the space between the atoms has to be divided according to the location of the atoms relative to each other. Various methods, based on Voronoi's idea of approximating the atomic space by polyhedra, have been proposed for this purpose. Comparing procedures concerned with the allocation of space among distinct atoms, we observe different partitionings of space, with deviations of more than 100% for particular atoms. Furthermore, we find that the separating planes of different Voronoi procedures do not meet the intersection circles of covalently linked atoms. This leads to a misallocation of space of up to 7% for atom pairs that largely differ in atomic size (e.g., C—H). Several algorithms are negatively affected by small unallocated polyhedra (''vertex error''). These effects are cumulative for a small protein up to a loss of some 60 $Å^3$ of total volume, which would correspond to the deletion of one complete residue. To overcome these errors, instead of using dividing planes between the atoms, we use curved surfaces, defined as the set of those geometrical loci with equal orthogonal distance to the surfaces of the van der Waals spheres under consideration. The proposed dividing surface meets not only the intersection circle of the two van der Waals spheres but also the intersection circle of the two spheres enlarged by an arbitrary value (e.g., radius of water). This hyperbolic surface enveloping the Voronoi cell can be easily constructed and offers the following advantages: no misallocation of volume for atoms of different size, no vertex error, geometrically reasonable allocation of the volume among atoms, avoidance of discontinuities between neighboring atoms, and improved applicability to water-accessible protein surfaces. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18**: 1113–1123, 1997

Keywords: Voronoi volumes; packing of atoms; density; proteins

Correspondence to: Dr. Cornelius Frömmel; e-mail: froemmel@comcom.rz.charite.hu-berlin.de

CCC 0192-8651 / 97 / 091113-11

Introduction

Generally, the atomic packing of proteins can be considered as nearly perfect.[1] In the core of a protein molecule, each atom is surrounded by a set of other atoms that contact their van der Waals surfaces. The resulting mean density (0.70−0.78) is close to that of organic crystals showing only local density[2] deviations. To achieve better comparisons of the function-relevant parts of the molecules, emphasis has been placed upon accurate volume calculations.[3,4] Furthermore, precise volume computation is required for predictions about thermostability of cavity-filling mutants.[5]

Algorithms for analytical computation of molecular volume,[6] solvent-excluded volume,[7] or ''molecular skin''[3] are time consuming and have to take a large number of singularities into account, where approximations are indispensable.

A common method for calculating volume and studying packing of spheres was originally developed by Voronoi[8] and was later applied to proteins by Richards.[1] Since then it has been used successfully as a standard method for the calculation of volumes of protein constituents, the description of protein motions, and the analysis of cavities in proteins.[2,9−17] The method has also been used in the analysis of computational molecular simulations.[18,19]

By Voronoi's procedure,[8] which is derived for a collection of points, the entire space is partitioned among a collection of equal-sized atoms. Each atom is surrounded by a polyhedron with its volume assigned to the atom. The faces of Voronoi polyhedra are derived from the separating planes perpendicular to the interatomic vectors, whereas the edges of the polyhedra are formed by the intersection of these planes. The Voronoi procedure requires the location of all atomic neighbors and a definition of the division points on the interatomic vectors. This can easily be done with molecules consisting of atoms with equal radii. In proteins, however, there are atoms showing different radii between 1.2 Å (hydrogen) and 1.8 Å (sulfur) and regions with varying local density.[2] The problem of nonassigned space is particularly pronounced for those atoms of proteins neighboring cavities: their neighbors are probably water molecules, which are often not localized in crystal structures.[17] In summary, modifications of Voronoi's original method are necessary to account for the peculiarities of protein structures.

Our primary aim in this article is to assess avoidable errors in various procedures used to calculate atomic volumes. The second goal is to derive an algorithm to estimate the dividing surface between the van der Waals surfaces of atoms. This separating surface is defined using a smooth 3-dimensional (3-D) function, which is dependent on the distance and radii of the neighboring atoms. The enveloped space is called the ''Voronoi cell.''

Methods

The total van der Waals volume is the space inside all of the van der Waals spheres. The van der Waals volume of an atom is the space inside the van der Waals sphere of the atom, cut by the separating surface between covalently linked atoms. The solvent excluded volume is the space that is not accessible to any center of the solvent spheres (rolling a solvent probe sphere over the protein surface).[7] This space is inside the spheres around the atoms with enlarged radii: the van der Waals radius of each atom is expanded by the radius of the solvent (see Fig. 1).

We used analytical as well as numerical calculations to compare the common algorithms for volume partitioning.

Calculations of the local atomic density[12] require the partitioning of the volume around the center of the atoms into two parts: the van der Waals volume of an atom itself and the volume assigned to this atom but outside the van der Waals sphere. For the protein interior the local atomic density can be calculated as the quotient of the van der Waals volume and the Voronoi (polyeder or cell) volume of the particular atom. To use this concept for atoms around packing defects and cavities, we considered only the solvent excluded volume in our calculations. Hence, the local atomic density is computed as the quotient of the van der Waals volume and the Voronoi cell volume limited to the solvent excluded sphere. This approach is necessary because up to 3.0 $Å^3$ cavities per residue are found in crystal structures and often no water molecules were localized. Not considering this effect causes the mean densities to be up to 2.0% lower than realistic values.[17]

The atomic coordinates of bovine pancreas trypsin inhibitor (PDB, code 5PTI) were selected

for comparing the known methods applied to proteins. Numerical computations were performed using a cubic lattice with a grid distance of 0.05 Å. The coordinate set included H atoms, allowing precise volume calculations for each type of atom. Different sets of atomic radii are included in the analysis.

For analytical estimations of volume, a simple two-atomic arrangement was selected from a well-refined structure of PDB.[20] Atoms Cβ and Hβ2 of the side chain of proline 2 were chosen from the high resolution structure of PDB (code 5PTI, resolution 1.0 Å). The total volume of the linked atoms can be easily computed from atomic radii (1.7 and 1.2 Å) and distance (1.05 Å), i.e., as the sum of the two spheres minus the two caps. Thus, an analytical solution for different positioning of the separating plane, sphere, or hyperboloid is possible.

In the original Voronoi procedure,[8] all atoms are considered as spheres with equal radius and each separating plane is positioned midway between two atoms (see Fig. 1A). However, atoms have different sizes. Therefore, several authors proposed a localization of the dividing plane according to atomic radii.

In the Richards method[1] the ratio of the distances between the atoms and the plane equals the ratio of the atomic radii (see Fig. 1B).

In the radical plane method Gellatly and Finney[21] positioned the plane in the intersection circle of the atoms to avoid vertex error (see Fig. 1C).

Gerstein et al.[19] first introduced nonplanar boundaries, called ''Gerstein spheres,'' between atoms, such that the ratio of distances to the centers of the neighboring atoms is constant (see Fig. 1D).

In the proposed Voronoi cell method, the separation surface consists of those loci with equal orthogonal distances to both van der Waals spheres (see Fig. 1E). The construction of the division point on the interatomic vector can be carried out by shrinkage (or expansion) of both atoms by a common δ until their spheres exactly touch. For its construction, see Figure 2. The resulting surface is part of a hyperboloid, as subsequently shown for the 3-D case.

Considering two atoms at a distance of d, with the centers at $p = (0, 0, 0)$ and $q = (d, 0, 0)$, and with radii $r1$ and $r2$ for any point of the surface, the following equations are valid: $D1 = r1 + \delta$ and $D2 = r2 + \delta$;

$$\delta = \|x - p\| - r1 = \|x - q\| - r2, \qquad (1)$$

thereby

$$(x^2 + y^2 + z^2)^{0.5} - r1 = \left((x - d)^2 + y^2 + z^2\right)^{0.5} - r2. \qquad (2)$$

Equation (2) is squared:

$$x^2 + y^2 + z^2 + \Delta r^2 - 2\Delta r(x^2 + y^2 + z^2)^{0.5} = x^2 - 2dx + d^2 + y^2 + z^2, \qquad (3)$$

where the difference of the radii is assumed to be nonnegative.

Therefore,

$$\Delta r^2 + 2dx - d^2 = 2\Delta r(x^2 + y^2 + z^2)^{0.5} \geq 0. \qquad (4)$$

Squaring eq. (4) and dividing by $4(d^2 - \Delta r^2)$ we get

$$\left(x - \frac{d}{2}\right)^2 - \frac{\Delta r^2}{d^2 - \Delta r^2}(y^2 + z^2) = \frac{\Delta r^2}{4}, \qquad (5)$$

where $x \geq d/2 - \Delta r^2/2d$. The equation is only valid for d larger than Δr, otherwise the larger atom contains the smaller one.

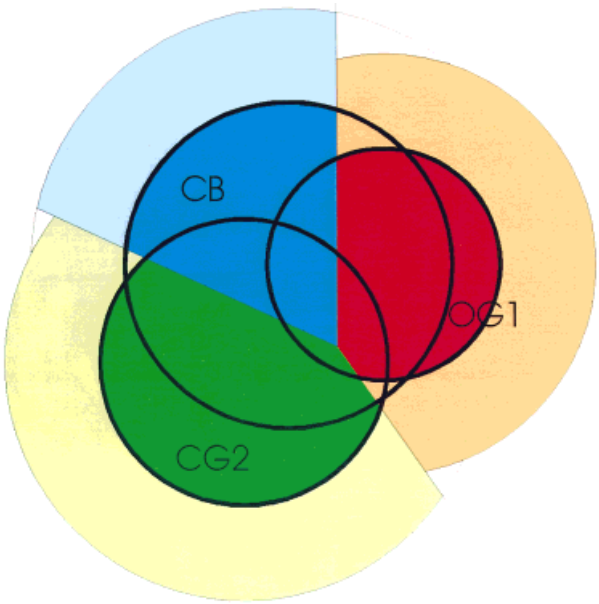
The general equation for hyperbolic surfaces is $ax^2 + by^2 + cz^2 = \text{const}$, where the coefficients a, b, and c are unequal to 0 and have different signs. The general equation for a rotation hyperbolic surface is $ax^2 + b(y^2 + z^2) = \text{const}$, which is fulfilled by eq. (5).

Results and Discussion

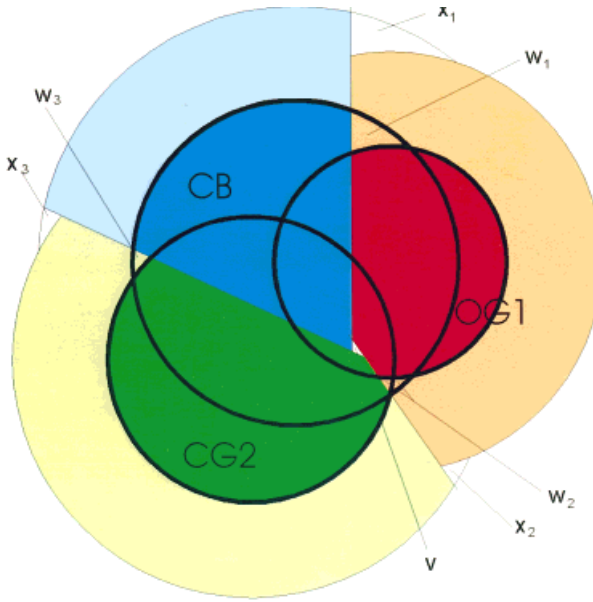
The results of the different methods are discussed for two applications: simple two-atomic ensemble and complete protein.

For two bonded atoms, the following criteria were used: Is the sum of their calculated van der Waals volumes equal to the total van der Waals volume? Is the assignment of the volume to the atoms reasonable? Are there overlapping parts of two atoms that should be part of the first atom but are assigned to the second (for instance, the center of the first atom)? Are there parts of one atom outside of the second atom that are assigned to the second (underestimation of the first atom)? The same questions are discussed for the solvent excluded volume.

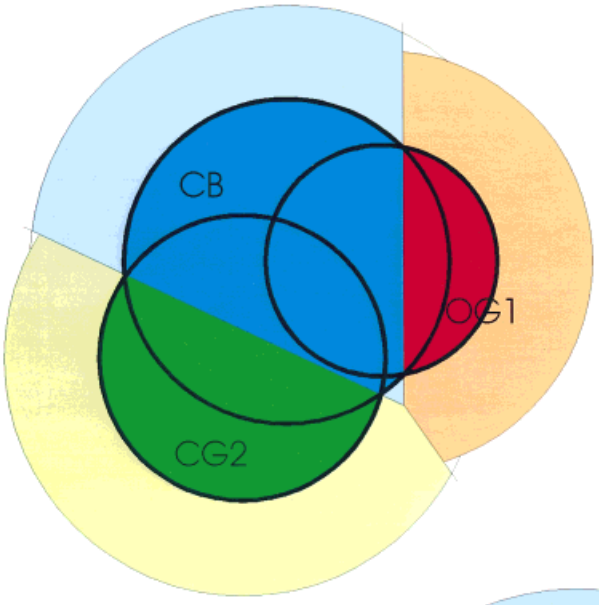
In the case of the protein, we address the following points: The application of the different procedures for volume calculations of proteins is ex-


CB
OG1
CG2

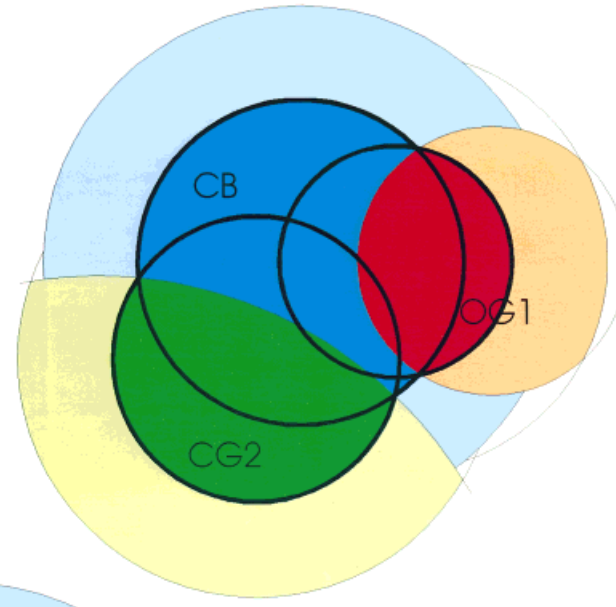
(a)


x1
w1
w3
x3
CB
OG1
CG2
w2
x2
v

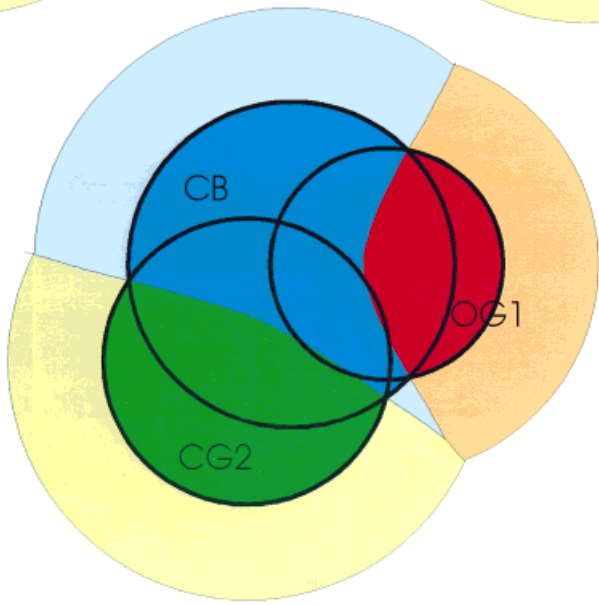
(b)


CB
OG1
CG2

(c)


CB
OG1
CG2

(d)


CB
OG1
CG2

(e)

amined in detail: the volume of different types of atoms is investigated. The influence of different sets of radii is also discussed. The volume and density calculations for proteins were carried out using a grid method. Gogonea and Osawa[22] compared the results of evaluations using grid methods with results from analytical algorithms. Numerical deviations between 1 and 2.4% were found for grids with grid values of 0.1 and 0.01 Å.[23,24] We estimated the volume of the protein using a cubic lattice with a grid distance of 0.05 Å. The final point addressed is the influence of the proposed separating surface on the computational effort of numerical and analytical volume calculations.

VORONOI POLYHEDRA

The most straightforward method is the original Voronoi method.[8] The algorithm defines intersection planes between atoms independently of the size of the participating atoms. Therefore, misallocations occur for intersecting atoms with different radii. The larger part of the overlapping volume is assigned to the smaller atom. Even the volume that is outside the small atom but inside the large atom is assigned to the small atom. A consequence of these characteristics is a loss of total van der Waals volume: the sum of the calculated van der Waals volumes is actually smaller than the total van der Waals volume (up to 7%).

Moreover, positive deviations of more than 100% in the volume of small atoms are found compared with the methods examined here, and the volume of larger atoms shows negative deviations of up to 30% (see Table I).

Volume or density calculations at the protein surface require supposition of a region occupied by solvent. The solvent accessible surface is defined as the set of loci of the center of a probe rolling over the protein surface.[25] The radius of the probe is usually constant and set to 1.4 Å, which is problematic and is discussed in detail by Gerstein et al.[19] To evaluate the applicability of various approaches of volume and density calculations for protein surfaces, we included calculations with all atom radii expanded by 1.4 Å (Table I). Using the original Voronoi method to compute this solvent excluded volume, the effects of misallocation of volume are even increased. For small atoms positive deviations of the solvent excluded volume occur (48 $Å^3$ instead of 13 $Å^3$; see Table I) and for large atoms negative differences appear (78 $Å^3$ instead of 113 $Å^3$; see Table I). A loss of total solvent excluded volume for two bonded atoms of up to 5% can be stated (see Table I).

Calculating the total volume of a protein, the deviation from correct values as described above occurs especially at the surface of the protein, in regions of smaller local density and near atoms with highly differing radii. Similar to the two-atomic ensemble, up to 7% of the total van der Waals volume of the protein still vanishes. Considering different types of atoms, the assigned volume differs to a high degree: smaller atoms like hydrogen are overestimated relative to the radical plane values up to 20% for the van der Waals volume and up to 15% for the solvent excluded volume. Larger atoms (carbon) are underestimated by about 20% of the van der Waals volume and up to 30% of the solvent excluded volume (see Table II) compared with the radical plane method.

FIGURE 1. Comparison of five methods for allocation of space among three atoms of threonine. The assigned volume of the van der Waals spheres are colored red, green, and blue. These spheres enlarged by the radius of the solvent (usually 1.4 Å) envelope the solvent excluded volume. The assigned volume of the corresponding solvent excluded spheres is colored with pale red, green, and blue colors. (A) In the original Voronoi procedure,[8] all atoms are considered as spheres with equal radius, and each separating plane is positioned midway between two atoms at distance D: $D1 = D/2$. (B) Richards method.[1] For covalently linked atoms, the distance between atom 1 and the plane is $D1 = D \cdot r1/(r1 + r2)$. The vertex error is indicated as (v). Errors occurring in various approaches are marked by w_i and x_i. The letter w designates loss of total volume resulting from the following constellation: the volume is located inside Cβ and outside Oγ1. Nevertheless, it belongs to the expanded volume of Oγ1 because it lies beyond the plane between Cβ and Oγ1. The letter x describes similar circumstances for the expanded atoms. (C) In the radical plane method Gellatly and Finney[21] positioned the plane to avoid vertex error, but the partition is not that reasonable: the center of the oxygen atom belongs to the carbon atom: $D1 = (D^2 + r1^2 - r2^2)/2D$. (D) Gerstein spheres.[19] These authors first introduced nonplanar boundaries between atoms such that the ratio of distances to the centers of the neighboring atoms was constant. The center of the dividing sphere is $C = D/(r1^2/r2^2 - 1)$ and the radius is $R = Cr1/r2$. (E) Voronoi cell. The partition area is positioned such that the distance to the spheres of the atoms of interest is the same.

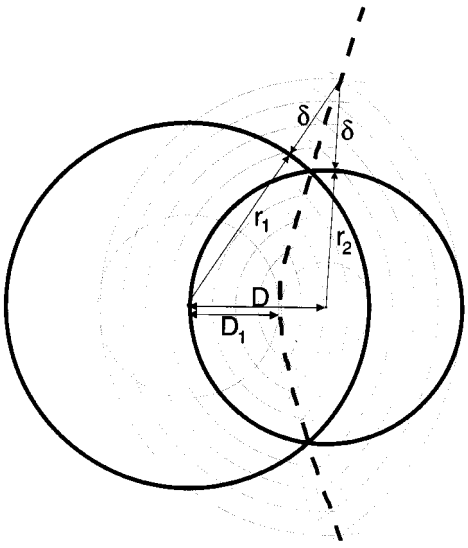


FIGURE 2. Construction scheme for the surface of the Voronoi cell around atoms. The separation surface is defined as the set of those loci with equal orthogonal distances to both van der Waals spheres. The construction of the division point on the interatomic vector can be carried out by shrinkage (or expansion) of both atoms by a common δ until their spheres exactly touch. With D_A and D_B distances of an arbitrary point P of the separating surface to the centers of the atoms, we assumed $D_A = r1 + \delta$ and $D_B = r2 + \delta$. Negative values of δ draw the surface within both spheres and $\delta = 0$ refers to the intersection circle. The resulting surface is part of a hyperboloid as demonstrated previously.

With regard to the original Voronoi method and its applicability for atomic density calculations, we attained the following results: the method shows the largest deviations of van der Waals volume and solvent excluded volume compared with other methods. The deviations in the density of the whole protein, however, are smaller (-10%) because the errors for the van der Waals volume and the solvent excluded volume partly compensate for each other. Generally, the relative error in the computation of the van der Waals volume is larger than that for the solvent excluded volume. Therefore, the estimated density for the interior protein is smaller than the density calculated by other methods (see Table III).

When considering the local density of different types of atoms, we notice an overestimation of the density of larger atoms and an underestimation of the density around smaller atoms. The reason for this observation is a larger deviation in the assignment of the solvent excluded volume than of the van der Waals volume. The deviations vary between -20% for hydrogen and $+15\%$ for carbon, oxygen, and sulfur.

For proteins the five methods were compared using the atomic radii from Bondi.[26] To assess the influence of the set of atomic radii, the evaluations were carried out using those from Gerstein et al.[19] As expected, the deviations described above were

TABLE I.
Comparison of Different Methods for Allocation of van der Waals Volume (V_{vdW}) and Solvent Excluded Volume ($V_{+1.4Å}$) of Two Covalently Linked Atoms.

Atom[a]	Original Voronoi[8] (Fig. 1A)	Richards[1] (Fig. 1B)	Radical Plane[21] (Fig. 1C)	Gerstein Sphere[19] (Fig. 1D)	Voronoi Cell[b] (Fig. 1E)
			V_{vdW} (Å^3)		
CB	14.91	15.63	19.44	18.00	18.62
HB2	5.84	5.50	2.88	4.32	3.70
Total	20.75	21.13	22.32[c]	22.32[c]	22.32[c]
			$V_{+1.4Å}$ (Å^3)		
CB	78.09	80.73	97.21	112.71	104.11
HB2	47.81	45.95	33.30	13.43	27.87
Total	125.90	126.68	130.51	126.14	131.98[c]

The two atoms are side chain atoms of proline 2 of bovine pancreas trypsin inhibitor (PDB code 5PTI). All values were calculated analytically. The volume calculation for caps of a sphere with radius r was performed using $V = 1/3\pi \cdot h^2(3r - h)$, with h being the height of the cap. The volume of a cap of a rotation hyperboloid $x^2 - b(y^2 + z^2) = q^2$ (with characteristic constants $b, q > 0$) can be calculated using $V = 1/3\pi \cdot h^2(3q + h)/b$ with h being the height of the caps.

[a] The heading CB characterizes the Cβ atom with radius 1.7 Å; HB2 marks the hydrogen atom with radius 1.2 Å.[26]

[b] The Voronoi cell method described in this article.

[c] The marked values are identical to the analytically calculated total volume of both spheres.

TABLE II.
Comparison of Different Methods for Allocation of van der Waals Volume (V_{vdW}) and Solvent Excluded Volume ($V_{+1.4Å}$) Considering All Atoms Including H Atoms in Bovine Pancreas Trypsin Inhibitor (PDB code 5PTI).

Type of Atoms	Original Voronoi[8]	Richards[1]	Radical Plane[21]	Gerstein Sphere[19]	Voronoi Cell[a]
			V_{vdW} (Å^3)		
Hydrogen	2698.9	2639.1	2196.3	2392.6	2239.1
Carbon	1688.5	1836.4	2329.8	2239.6	2370.8
Oxygen	538.7	587.7	715.4	696.7	733.2
Nitrogen	418.4	416.0	364.5	399.0	384.0
Sulfur	30.1	32.8	48.1	45.1	49.5
Total	5380.3	5517.5	5657.3	5777.4	5780.6
			$V_{+1.4Å}$ (Å^3)		
Hydrogen	7926.2	7539.9	6903.2	5641.2	6471.2
Carbon	2374.9	2726.6	3312.3	4127.0	3705.9
Oxygen	681.7	790.1	953.5	1153.1	1080.3
Nitrogen	1325.9	1276.6	1166.6	985.4	1115.0
Sulfur	30.1	33.0	51.0	62.8	57.8
Total	12373.5	12400.0	12416.8	11985.2	12457.4

A cubic lattice with a grid distance of 0.05 Å was selected for all five methods using the following radii: C 1.7 Å, N 1.55 Å, O 1.52 Å, S 1.8 Å, H 1.2 Å.[26]

somewhat smaller using this set because of the smaller differences between radii.

RICHARDS' METHOD

At first Richards[1] criticized the localization of the dividing plane independent on the radii of the joined atoms. As a first approximation he derived a method positioning the plane according to the ratio of the radii of the participating atoms. He noticed the neglect of tiny tetrahedrons near each polyhedron vertex (''vertex error''). But this approximation provided a chemically rational division and was widely used later. Gerstein et al.[19] noticed a vertex error of about 46.0 Å^3 for the volume of the examined protein (PDB, code 4PTI).

TABLE III.
Comparison of Local Density for Interior Atoms of Bovine Pancreas Trypsin Inhibitor (PDB code 5PTI) Computed by Different Methods.

Type of Atoms	Original Voronoi[8]	Richards[1]	Radical Plane[21]	Gerstein Sphere[19]	Voronoi Cell[a]
Hydrogen	0.553	0.597	0.564	0.705	0.643
Carbon	0.878	0.829	0.834	0.756	0.794
Oxygen	0.937	0.877	0.882	0.769	0.816
Nitrogen	0.660	0.675	0.675	0.738	0.703
Sulfur	0.998	0.991	0.956	0.836	0.903
Cα Asn43[b]	1.000	1.000	0.998	0.901	0.947
All types[c]	0.654	0.682	0.688	0.733	0.719

[a] The Voronoi cell method described in this article.
[b] For this particular atom the local packing density evaluated by the different methods is given as an example. The values are computed without consideration of included solvent molecules. The atom is situated near a water-filled cavity but is slightly accessible to this water.
[c] Mean packing density of all interior atoms.

Nevertheless, with this method the specialties of atomic ensembles were implemented to the original Voronoi procedure.

For the two-atomic ensemble the method results in about a 5% loss of total van der Waals and solvent excluded volume, respectively (see Table I). The assignment to particular atoms shows a trend of stronger weight on large atoms compared with the original Voronoi procedure.

Evaluating the total volume of a protein, the observed deviations are smaller than those in the two-atomic case. A loss of about 4% of the van der Waals volume of the protein was still observed (see Table II). The estimated solvent excluded volume was closer to the correct value (−1%).

Similar to the original Voronoi procedure, the total evaluated density of a protein was underestimated (up to 6%). For particular types of atoms the mean local density deviated about 10% from our results (see Table III).

RADICAL PLANE

A principal advantage of the radical plane method[21] separation scheme is the passage of the dividing plane through the intersection circle of both atoms. Thus, the estimation of the total volume of two linked atoms is free of systematic errors. Due to this advantage, the method is frequently used to estimate the van der Waals volume of proteins. However, the partition of the volume is not always chemically reasonable. For bonded atoms (distance about 1.4 Å), for example, the center of an atom is not part of the volume assigned to the atom itself if the difference of the radii is at least 0.6 Å (see Fig. 1C, atom OG1). Deviations of the total volume appear only for the solvent excluded volume and are small (about 1%).

The points mentioned above also remain valid for estimations of volume and density in whole proteins (see Table II). However, the density of an entire protein can still be calculated with sufficient precision using this method. Differences between the local density estimated by the radical plane method and the Voronoi cell procedure occur especially in the case of bonded atoms with differing radii, even if packing defects are observed nearby (see Table III).

GERSTEIN SPHERES

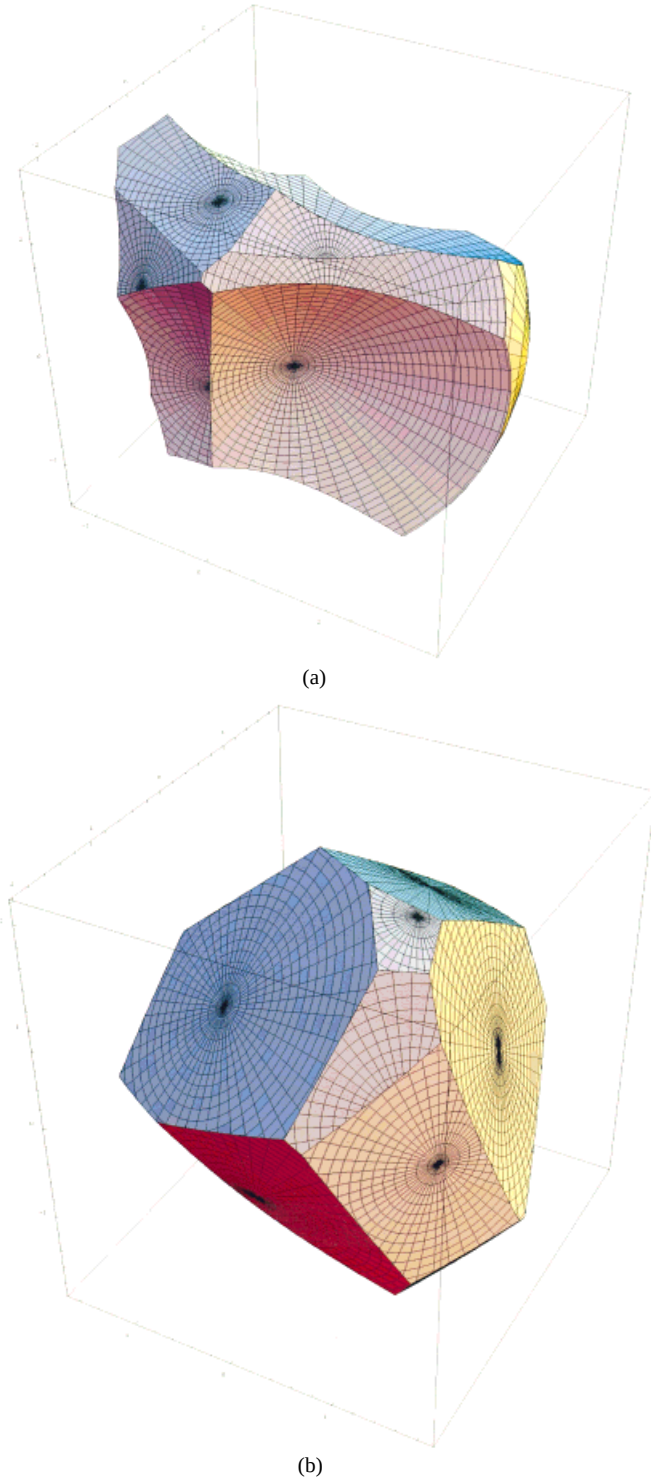
The partition surface between two or more atoms constructed by the method of Gerstein et al.[19] meets the intersection circle of overlapping areas as the radical plane method does. Therefore, the total volume of bonded atoms is correctly estimated. The method accomplishes a geometrically reasonable partition of the volume among atoms by use of a curved surface as a border. But the method results in up to a 4.5% loss of total solvent excluded volume for the two-atomic case and a 4.0% loss for the entire protein. For energy optimizations precise volume calculations are necessary. An erroneous estimation like that of the solvent excluded volume of up to 400 $Å^3$ (see Table II) can lead to a deviation in energy calculations of more than 40 kcal/mol.[22]

For the solvent excluded volume of small bonded atoms negative deviations of up to 50% in comparison with the method proposed here occur. Therefore, larger deviations in local densities will occur. The local density assigned to smaller bonded atoms will appear higher than expected. Determining the total density of a protein using Gerstein spheres, the computed value is somewhat higher than the precise value because of the loss of solvent excluded volume (see Table III).

VORONOI CELL

The Voronoi cell method proposed here unifies the advantages of earlier approaches by keeping the spirit of the geometrically rational partitioning implied by the Richards method that avoids vertex error and meets the intersection circle between the atoms like the radical plane method and using nonplanar boundaries like Gerstein et al.[19] Voronoi cells show a complex 3-D shape. To give an impression of the variety of shapes of Voronoi cells, we selected two diverse examples. Voronoi cells with concave interfaces occur for large atoms with small neighbors (see Fig. 3A). Conversely, convex interfaces build up Voronoi cells for small atoms with larger neighbors. To illustrate the first case, we show a phosphorus atom (radius 2.2 Å) from a phosphate ion (radius of oxygen 1.4 Å) located at the protein surface of the pancreas trypsin inhibitor (PDB code 5PTI) (see Fig. 3A). In the second case, a water molecule (radius 1.4 Å) inside a protein cavity of lysozyme (PDB code 135L) is shown (see Fig. 3B).

An example for deviations of the local packing density computed by different methods is given in Table III. The examined Cα atom is situated near a water-filled cavity. It is slightly accessible to this water molecule, which is widely neglected by three methods (density = 1.0).



(a)

(b)

FIGURE 3. Three-dimensional representation of Voronoi cells for the solvent excluded volume. For visualization a ''spider web'' is placed on every dividing hyperboloid with its center on the interatomic vector between the atom under consideration and its neighbors. In general, for solvent-inaccessible atoms of proteins, about 20 neighbors have to be considered. (A) Phosphorus atom (radius 2.2 Å) from a phosphate ion (radius of oxygen 1.4 Å) at the protein surface of the pancreas trypsin inhibitor (PDB code 5PTI). Concave interfaces dominate because neighboring atoms are smaller than the phosphorus. The Voronoi cell for the solvent excluded volume of the phosphorus was cut at a distance of 3.6 Å (atom radius + 1.4 Å). (B) A water molecule inside a protein cavity of lysozyme (PDB code 135L). Hyperboloids of Voronoi cells for small atoms with larger neighbors are convex. In this case, neighboring water molecules result in dividing planes.

There is neither loss of total volume nor of total solvent excluded volume because the partition surface meets the intersection circle of the van der Waals spheres as the intersection circle of the solvent excluded spheres. No vertex error can occur.[27] Consequently, the density of the entire protein is evaluated precisely. The assignment of the volume to particular atoms and resulting local densities appear reasonable. The advantages of the method are fully independent of the choice of atomic radii or solvent radius.

Computational Effort

In principal, there are numerical and analytical procedures possible for all five methods discussed here. The effort of grid methods increases generally by $O(N^3)$ with decreasing grid distance or by $O(N^2)$ for Monte Carlo sampling.[19] When using the proposed Voronoi cell procedure, the grid points can be tested in a similar way as for the original Voronoi procedure: instead of checking some linear inequalities, some quadratic inequalities [see eq. (5)] have to be examined, which increases the effort by less than 10%. For the considered protein the density calculation takes a few minutes on an IBM PC (grid distance 0.05 Å).

Analytical methods based on the Gauss–Bonnet theorem[28] integrate over the area of the intersecting surface between a growing sphere and the corresponding polyhedron. A large number of singularities have to be considered because during the growth process of the sphere, the case of touching circles permanently occurs, as well as the case of three or more circles intersecting in one point. This problem is increased by the computational error. A further complication occurs for enlarged spheres when evaluating the solvent excluded volume because the number of intersections rises dramatically (by 2 orders of magnitude). Using analytical methods for the volume computation, the expenditure of computing time is nearly the same for all types of separating surfaces. The Voronoi cell approach can be used not only in a grid method, but is also suitable for the analytical evaluation of volumes similar to the method proposed by Gogonea and Osawa.[22] The implementation of the algorithm is in progress. An exclusive advantage of the Voronoi cell will be that the van der Waals volume and the solvent excluded volume can be evaluated simultaneously. This results from the intersection of the dividing surface through both the intersection circles of van der Waals spheres and those of enlarged spheres.

Conclusions

For all partitioning schemes derived from polyhedra, the allocated volume depends strongly on the location of the interatomic plane. A detailed analysis of known methods (Fig. 1) shows that the results differ by more than 100% for small atoms (Table I). The method proposed here avoids the systematic loss of total volume of proteins (up to 7%) and is therefore additive for the van der Waals volume as well as for the solvent excluded volume. This permits the calculation of molecular volume simply by addition of the atomic values.

Appendix: Proof of Absence of Vertex Error for Different Methods for Allocation of Space

Generally, the separating surface between two atoms x_i and x_j can be defined implicitly by a function $f(x, x_i, x_j) = 0$. All points x fulfilling the equation belong to the separating surface between the two atoms: points with $f(x, x_i, x_j) < 0$ are assigned to one atom; points with positive $f(x, x_i, x_j)$ are assigned to the other. The function that defines the surface is not definite: different functions can define the same surface. If there is a function that can be written in the form

$$f(x, x_i, x_j) = g(x, x_i) - g(x, x_j), \quad \text{(A.1)}$$

meaning that the variables x_i and x_j can be separated, no vertex error can appear. Possible functions for the different methods are given below. The condition (A.1) can be easily checked by proofing the following condition: Is the difference $f(x, x_i, x_j) - f(x, x_0, x_j)$ independent from variable x_j? This becomes clear for a decomposition of type (A.1):

$$\begin{aligned} f(x, x_i, x_j) - f(x, x_0, x_j) &= g(x, x_i) - g(x, x_j) - g(x, x_0) + g(x, x_j) \\ &= g(x, x_i) - g(x, x_0). \end{aligned}$$

The same has to be proofed for $f(x, x_i, x_j) - f(x, x_i, x_0)$. The condition needs to be fulfilled only for an arbitrary function defining the surface. It is a sufficient condition for the absence of vertex

error. The vertex error occurs if there are points x for which $f(x, x_i, x_j) > 0$ (meaning that x is not assigned to x_i but possibly assigned to x_j), $f(x, x_j, x_k) > 0$ (x is not assigned to x_j, possibly assigned to x_k), and $f(x, x_k, x_i) > 0$ (x is not assigned to x_k, possibly assigned to x_i). In that case the point x cannot be assigned to any atom. A circular reasoning like this is not possible for functions of type (A.1), because $f(x, x_i, x_j) > 0$ and $f(x, x_j, x_k) > 0$ implies $f(x, x_k, x_i) < 0$.

The reason for that is

$$f(x, x_k, x_i) = -f(x, x_i, x_k) = -\big(f(x, x_i, x_j) + f(x, x_j, x_k)\big) < 0.$$

It can be easily seen that the three separating surfaces created by three atoms meet together on a single curve, and the six surfaces formed by four atoms meet together at a single point when the surface is created by functions of type (A.1).

Possible functions of type (A.1) for the different methods of volume allocation. The planes of original Voronoi can be defined using

$$g(x, x_i) = d(x, x_i),$$

with $d(x, x_i)$ being the Euclidean distance of the points x and x_i.

The planes of the radical plane method result from

$$g(x, x_i) = d^2(x, x_i) - r_i^2.$$

The spheres of Gerstein can be defined using

$$g(x, x_i) = d(x, x_i)/r_i,$$

with radius r_i of the atom x_i.

The proposed method is created by the function

$$g(x, x_i) = d(x, x_i) - r_i.$$

References

1. F. M. Richards, *J. Mol. Biol.*, **82**, 1 (1974).
2. F. M. Richards, *Annu. Rev. Biophys. Bioeng.*, **6**, 151 (1977).
3. B. B. Masek, A. Merchant, and J. B. Matthew, *J. Med. Chem.*, **36**, 1230 (1993).
4. T. D. J. Perkins, J. E. J. Mill, and P. M. Dean, *J. CAMD*, **9**, 479 (1995).
5. K. Ishikawa, H. Nakamura, K. Morikawa, and S. Kanaya, *Biochemistry*, **32**, 6171 (1993).
6. M. L. Connolly, *J. Am. Chem. Soc.*, **107**, 1118 (1985).
7. V. Gogonea and E. Osawa, *J. Comput. Chem.*, **16**, 817 (1995).
8. G. F. Voronoi, *J. Reine Angew. Math.*, **134**, 198 (1908).
9. C. Chothia, *Nature*, **248**, 338 (1974).
10. Y. Harpaz, M. Gerstein, and C. Chothia, *Structure*, **2**, 641 (1994).
11. J. Janin and C. Chothia, *J. Biol. Chem.*, **265**, 16027 (1990).
12. F. M. Richards, *Carlsberg Res. Commun.*, **44**, 47 (1979).
13. F. M. Richards, *Methods Enzymol.*, **115**, 440 (1985).
14. M. Gerstein, A. M. Lesk, E. N. Baker, B. Anderson, G. Norris, and C. Chothia, *J. Mol. Biol.*, **234**, 357 (1993).
15. J. L. Finney, *J. Mol. Biol.*, **96**, 721 (1975).
16. J. L. Finney, B. J. Gellatly, I. C. Golton, and J. Goodfellow, *Biophys. J.*, **32**, 17 (1980).
17. S. J. Hubbard and P. Argos, *Curr. Opin. Biotechnol.*, **6**, 375 (1995).
18. J. P. Shih, S. Y. Sheu, and C. Y. Mou, *J. Chem. Phys.*, **100**, 2202 (1994).
19. M. Gerstein, J. Tsai, and M. Levitt, *J. Mol. Biol.*, **249**, 955 (1995).
20. F. Bernstein, T. Koetzle, G. Williams, E. Meyer, Jr., M. Brice, J. Rodgers, O. Kennard, T. Shimanuchi, and M. Tasumi, *J. Mol. Biol.*, **112**, 535 (1977).
21. B. J. Gellatly and J. L. Finney, *J. Mol. Biol.*, **161**, 305 (1982).
22. V. Gogonea and E. Osawa, *Supramol. Chem.*, **3**, 303 (1994).
23. A. Gavezzotti, *J. Am. Chem. Soc.*, **105**, 5220 (1983).
24. A. Y. Meyer, *J. Chem. Soc. Perkin Trans.*, **2**, 1161 (1985).
25. B. Lee and F. M. Richards, *J. Mol. Biol.*, **55**, 379 (1971).
26. A. Bondi, *Phys. Chem.*, **68**, 441 (1964).
27. F. Aurenhammer and H. Imai, *Lect. Notes Comput. Sci.*, **247**, 53 (1987).
28. M. P. Do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice Hall, Englewood Cliffs, NJ, 1976.